

# Picard iterations for diffusions on symmetric matrices

Carlos G. Pacheco\*

May 14, 2015

## Abstract

Matrix-valued stochastic processes have been of significant importance in areas such as physics, engineering and mathematical finance. One of the first models studied has been the so-called Wishart process, which is described as the solution of a stochastic differential equation in the space of matrices. In this paper we analyze natural extensions of this model, and prove the existence and uniqueness of the solution. We do this by carrying out a Picard iteration technique in the space of symmetric matrices. This approach takes into account the operator character of the matrices, which helps to corroborate how the Lipchitz conditions also arise naturally in this context.

## 2000 Mathematics Subject Classification:

*Keywords:* Matrix-valued diffusions, Lipschitz conditions, Picard iterations.

## 1 Introduction

Bru [2] introduced the so called Wishart process, which is specified by the following stochastic differential equation (SDE) valued in the space of symmetric  $d \times d$  matrices for certain values  $\alpha$  in the so-called Wallach set (i.e.  $\alpha \in \{1, 2, \dots, d-1\} \cup [d-1, \infty)$ ) and some initial condition  $X_0$ ,

$$dX_t = \sqrt{X_t}dB_t + dB_t^T\sqrt{X_t} + \alpha Idt,$$

where  $B_t$  is a  $d \times d$  matrix with each entry being a Brownian motion, all of them independent. If  $X_t$  belongs to the set of positive semidefinite symmetric matrices, one can properly define  $\sqrt{X_t}$  for each  $t \geq 0$ , as  $\sqrt{X_t} = U_t\sqrt{\Lambda_t}U_t^T$ , where  $U_t\Lambda_tU_t^T$  is the spectral decomposition of  $X_t$ ,

---

\*Departamento de Matematicas, CINVESTAV-IPN, A. Postal 14-740, Mexico D.F. 07000, MEXICO.  
Email: cpacheco@math.cinvestav.mx

and  $\sqrt{\Lambda_t}$  represents the diagonal matrix where the diagonal is given by the square roots of the eigenvalues increasingly ordered.

The work in [2] has motivated several studies and applications, let us mention for example [10, 13, 12]. Under certain conditions, Bru tells us that such an equation has a unique strong solution. To argue that, Bru [2] appeals first to the fact, taken from [15], that the square root is an analytic function in the space of symmetric positive matrices. Then, she refers to the result in Ikeda and Watanabe [8] about the unique existence of a solution of a vector-valued diffusion to conclude the uniqueness and existence of the solution. Here we propose taking a different route, where one needs some results on matrices.

A more general model is given by

$$dX_t = g(X_t)dB_t f(X_t) + f(X_t)dB_t^T g(X_t) + b(X_t)dt, \quad (1)$$

where  $g$ ,  $f$  and  $b$  are matrix valued functions acting on matrices. One example is when one takes  $\mathbb{R} \rightarrow \mathbb{R}$  functions and uses spectral decomposition to obtain matrix-valued functions; actually, we were motivated to study equation (1) after seeing this point of view in [5].

In this paper we propose using the Picard iteration method to establish the existence of the solution, as well as exploiting the operator character of the matrix to study the equation. We think that dealing directly (instead of passing through vector-valued diffusions) with the matrix-equation can be useful to obtain more insight into matrix diffusions. Thus, we hope that our results help to complement theory already developed in papers such as [1, 2, 3, 4, 6, 9, 13, 14, 16]. To carry out the proof, we develop a few results suited to handle matrix-equations and which are of independent interest.

One important issue that we are not studying here is the so-called *time of collision*; one might read more about this in [11, 12, 14, 17].

## 2 Preliminaries

Let  $\mathcal{S}_{d \times d}$  be the set of symmetric matrices, and  $\mathcal{S}_{d \times d}^+$  the positive semidefinite ones. Let  $B_t$  be a  $d \times d$  Brownian motion (i.e. a matrix filled with independent Brownian motions). We will focus on the following SDE valued in  $\mathcal{S}_{d \times d}$ :

$$dX_t = g(X_t)dB_t f(X_t) + f(X_t)dB_t^T g(X_t) + b(X_t)dt,$$

with initial condition  $X_0 \in \mathcal{S}_{d \times d}^+$ , and where  $g$ ,  $f$  and  $b$  are  $\mathcal{S}_{d \times d} \rightarrow \mathcal{S}_{d \times d}$  functions. However, we will be more interested in considering  $\mathbb{R} \rightarrow \mathbb{R}$  functions to construct a diffusion. In this case the following consideration is taken for  $\mathbb{R} \rightarrow \mathbb{R}$  functions  $g$ ,  $f$  and  $b$ . If  $A \in \mathcal{S}_{d \times d}$ , by  $g(A)$  we mean  $Hg(\Lambda)H^T$ , where  $H\Lambda H^T$  is the spectral decomposition of  $A$  and  $g(\Lambda)$  is the

diagonal matrix with the values  $g(\lambda_1), \dots, g(\lambda_d)$  and  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_d$  are the eigenvalues of  $A$  increasingly ordered. Under this framework it turns out that  $X_t$  is a symmetric matrix for all  $t$ . In [7], Chapter 6, there is a detailed study of functions acting on spaces of matrices, an idea which is extended in functional analysis to so-called *functional calculus* to define a function of an operator.

Our aim is to investigate the condition on the functions  $g$ ,  $f$  and  $b$ , under which previous equation has a unique strong solution. As expected, Lipschitz conditions will play a crucial role. Before we embark on this task, some useful results are in order.

**Definition 1** *For matrices  $A$  and  $B$ , by*

$$A \geq 0$$

*we mean  $A$  is positive semidefinite, that is  $x^T A x \geq 0$  for every vector  $x$ , and by*

$$A \leq B$$

*we mean  $B - A \geq 0$ .*

**Remark 2** *The following results will be useful (we used [19] as a general reference).*

*i) For symmetric matrices  $A, B$ ,*

$$(A + B)^2 \leq 2A^2 + 2B^2. \quad (2)$$

*ii) For a symmetric matrix  $A$  and unit vector  $x$ ,*

$$(x^T A x)^2 \leq x^T A^2 x. \quad (3)$$

*A proof of (3) can be obtained using the Cauchy-Schwarz inequality. Indeed*

$$(x^T A x)^2 = \langle x, A x \rangle^2 \leq \langle x, x \rangle \langle A x, A x \rangle = x^T A^T A x = x^T A^2 x.$$

Next, we prove an analogous result of the Cauchy inequality, which will be useful.

**Proposition 3** *Let  $\{A_t, t \geq 0\}$  in  $\mathcal{S}_{d \times d}$  with each entry being a continuous function. Then*

$$\left( x^T \int_0^t A_s ds x \right)^2 \leq t x^T \int_0^t A_s^2 ds x,$$

*for any unit vector  $x$ .*

**Proof.** First fix a unit vector  $x$ . Now consider an equidistant partition  $\{s_1, \dots, s_n\}$  of  $[0, t]$  and set  $A_i := A_{s_i}$  for  $i = 1, \dots, n$ . Let  $\Delta > 0$  be the partition size and define

$$F(u) := x^T \sum_{i=1}^n (\sqrt{\Delta} A_i - u \sqrt{\Delta} I)^2 x.$$

We have

$$\sum_{i=1}^n (\sqrt{\Delta} A_i - u \sqrt{\Delta} I)^2 = \underbrace{\Delta A_1^2 + \dots + \Delta A_n^2}_{\alpha} - 2u \underbrace{(\Delta A_1 + \dots + \Delta A_n)}_{\beta} + u^2 \underbrace{(\Delta + \dots + \Delta)}_t.$$

Since  $F(u) = x^T \alpha x - 2u x^T \beta x + u^2 x^T t x \geq 0$  for all  $u$ , then for the discriminant

$$(-2x^T \beta x)^2 - 4x^T \alpha x x^T t x \leq 0.$$

But  $x^T t x = t$ , hence

$$x^T (\Delta A_1^2 + \dots + \Delta A_n^2) x t \geq (x^T (\Delta A_1 + \dots + \Delta A_n) x)^2.$$

The result follows after taking the infinitesimal sum on both sides of the previous inequality.  $\blacksquare$

### 3 Existence and uniqueness

We will use the following criterion to establish the solubility of the stochastic equation.

**Definition 4** Consider a function  $g : \mathcal{S}_{d \times d} \rightarrow \mathcal{S}_{d \times d}$ . We say that  $g$  is **Lipschitz in matrix sense** if there exists a constant  $c > 0$  such that for any pair  $A_1, A_2 \in \mathcal{S}_{d \times d}$  and any unit vector  $x \in \mathbb{R}^d$  we have

$$x^T (g(A_1) - g(A_2))^2 x \leq c x^T (A_1 - A_2)^2 x. \quad (4)$$

The following two results will be useful for Theorem 7, their proofs are left in the Appendix, where it is properly defined what we mean by the matrix stochastic integral. Notice that the next proposition resembles an isometry property,

**Proposition 5** Let  $A_t$  and  $C_t$ ,  $t \geq 0$  be matrix-valued stochastic processes such that

$$\int_0^t A_s dB_s C_s$$

is well defined as an Itô stochastic integral. Then, for any pair of vectors  $x, y \in \mathbb{R}^d$

$$E \left[ y^T \left( \int_0^t A_s dB_s C_s \right)^2 x \right] = \int_0^t E \left[ x^T C_s^T C_s A_s A_s^T y \right] ds.$$

**Lemma 6** Let  $\tau > 0$  be fixed. For continuous adapted processes  $A_t$  and  $C_t$  in  $\mathcal{S}_{d \times d}$ , there exists  $\beta > 0$  such that

$$\begin{aligned} E \left[ x^T \left( \int_0^t A_s dB_s C_s + \int_0^t C_s dB_s^T A_s \right)^2 x \right] &\leq \beta \left| E \left[ x^T \left( \int_0^t A_s dB_s C_s \right)^2 x \right] \right| \\ &+ \beta \left| E \left[ x^T \left( \int_0^t C_s dB_s^T A_s \right)^2 x \right] \right|, \end{aligned}$$

for all  $t \in [0, \tau]$ .

**Theorem 7** Suppose that  $g$ ,  $f$  and  $b$  are  $\mathbb{R} \rightarrow \mathbb{R}$  bounded functions that satisfy the property of Definition 4. Then, the stochastic differential equation

$$X_t = X_0 + \int_0^t b(X_s) ds + \int_0^t g(X_s) dB_s f(X_s) + \int_0^t f(X_s) dB_s^T g(X_s) \quad (5)$$

has a unique strong solution in  $\mathcal{S}_{d \times d}$ .

The following proof follows the general structure of its vector analogue taken from [18].

**Proof.** Since  $g$ ,  $f$  and  $b$  are bounded, there is a constant  $c > 0$  such that for any symmetric matrix  $A$

$$g(A) < cI, \quad (6)$$

and the same for  $f$  and  $b$ .

**Uniqueness.** Let  $\tau > 0$  be fixed and consider  $t \in [0, \tau]$ . If  $X_t$  and  $Y_t$  are two solutions of the SDE (5), we want to prove that  $(\forall x \in \mathbb{R}^d) E [x^T (X_t - Y_t)^2 x] = 0$ .

Using inequality (2) we have

$$\begin{aligned} E [x^T (X_t - Y_t)^2 x] &= E \left[ x^T \left( \int_0^t (b(X_s) - b(Y_s)) ds + H_t(X, Y) \right)^2 x \right] \\ &\leq 2E \left[ x^T \left( \left( \int_0^t (b(X_s) - b(Y_s)) ds \right)^2 + (H_t(X, Y))^2 \right) x \right] \end{aligned}$$

where

$$H_t(X, Y) := \int_0^t g(X_s)dB_s f(X_s) + \int_0^t f(X_s)dB_s^T g(X_s) - \int_0^t g(Y_s)dB_s f(Y_s) - \int_0^t f(Y_s)dB_s^T g(Y_s)$$

With Proposition 3 and the Lipschitz condition (4) we have that

$$E \left[ x^T \left( \int_0^t (b(X_s) - b(Y_s))ds \right)^2 x \right] \leq \tau c \int_0^t E[x^T (X_s - Y_s)^2 x] ds. \quad (7)$$

Next we do the following for the other term,

$$\begin{aligned} H_t(X, Y) &= H_t(X, Y) \pm \int_0^t g(X_s)dB_s f(Y_s) \pm \int_0^t f(Y_s)dB_s^T g(X_s) \\ &= \underbrace{\int_0^t g(X_s)dB_s (f(X_s) - f(Y_s)) + \int_0^t (f(X_s) - f(Y_s))dB_s^T g(X_s)}_{H^{(1)}} \\ &\quad + \underbrace{\int_0^t (g(X_s) - g(Y_s))dB_s f(X_s) + \int_0^t f(X_s)dB_s^T (g(X_s) - g(Y_s))}_{H^{(2)}}. \end{aligned}$$

Notice that last expression is the sum of two symmetric matrices,  $H^{(1)}$  and  $H^{(2)}$ , thus, to analyze

$$E \left[ x^T (H_t(X, Y))^2 x \right],$$

we can apply inequality (2) to split the previous expression into two parts, one with  $H^{(1)}$  and the other with  $H^{(2)}$ . After that, we can apply Lemma 6 to each part with  $H^{(i)}$ , so that in the end we have split it into four terms. This means that  $E \left[ x^T (H_t(X, Y))^2 x \right]$  is less than or equal to the sum of four terms, each one of the form

$$E \left[ x^T \left( \int_0^t g(X_s)dB_s (f(X_s) - f(Y_s)) \right)^2 x \right].$$

Using Proposition 5, the Lipschitz (4) and the boundedness conditions (6), the following happens to each term

$$\begin{aligned} E \left[ x^T \left( \int_0^t g(X_s)dB_s (f(X_s) - f(Y_s)) \right)^2 x \right] &= \int_0^t E \left[ x^T (f(X_s) - f(Y_s))^2 g^2(X_s) x \right] ds \\ &\leq c_1 \int_0^t E[x^T (X_s - Y_s)^2 x] ds, \end{aligned}$$

for some finite constant  $c_1$ . All this, together with (7), ends up giving that

$$E [x^T (X_t - Y_t)^2 x] \leq c_\tau \int_0^t E [x^T (X_s - Y_s)^2 x] ds,$$

where  $c_\tau$  is a constant depending on  $\tau$ . An application of Gronwall's Lemma finishes this part of the proof, which is to say that  $E [x^T (X_t - Y_t)^2 x]$  is in fact zero for all unit vectors  $x$ .

**Existence.** The Picard iteration technique commands us to define

$$X_t^{(n)} := X_0 + \int_0^t b(X_s^{(n-1)}) ds + \int_0^t g(X_s^{(n-1)}) dB_s f(X_s^{(n-1)}) + \int_0^t f(X_s^{(n-1)}) dB_s^T g(X_s^{(n-1)}),$$

and  $X_t^{(0)} := X_0$  for all  $t \geq 0$ . We want to prove that there exists a stochastic process  $X_t$  valued in  $\mathcal{S}_{d \times d}$  such that

**i)**  $X_t^{(n)} \rightarrow X_t$  uniformly on  $t \in [0, \tau]$  and **ii)** that  $X_t$  satisfies the SDE (5).

First, in order to prove **i)**, with techniques already used in the Uniqueness part, i.e. inequality (2), Lemma 6, Proposition 5, as well as the boundedness condition (6), we have

$$\begin{aligned} & E[x^T (X_t^{(1)} - X_t^{(0)})^2 x] \\ &= E \left[ x^T \left( \int_0^t b(X_s^{(0)}) ds + \int_0^t g(X_s^{(0)}) dB_s f(X_s^{(0)}) + \int_0^t f(X_s^{(0)}) dB_s^T g(X_s^{(0)}) \right)^2 x \right] \\ &\leq 2t^2 E[x^T b^2(X_0)x] + 2\beta t^2 E[x^T f^2(X_0)g^2(X_0)x] + 2\beta t^2 E[x^T g^2(X_0)f^2(X_0)x] \leq c_\tau, \end{aligned} \quad (8)$$

for all  $t \in [0, \tau]$ , where  $c_\tau$  is a finite constant depending on  $\tau$ .

Now, using again (2), we have the following inequality

$$\begin{aligned} & E[x^T (X_t^{(n+1)} - X_t^{(n)})^2 x] \\ &= E \left[ x^T \left( \int_0^t (b(X_s^{(n)}) - b(X_s^{(n-1)})) ds + H_t(X^{(n)}, X^{(n-1)}) \right)^2 x \right] \\ &\leq 2E \left[ x^T \left( \int_0^t (b(X_s^{(n)}) - b(X_s^{(n-1)})) ds \right)^2 x \right] + 2E \left[ x^T (H_t(X^{(n)}, X^{(n-1)}))^2 x \right]. \end{aligned}$$

Let us analyze the last two terms in the left hand side of the previous display. With Proposition 3 and the Lipschitz condition (4) we obtain

$$E \left[ x^T \left( \int_0^t (b(X_s^{(n)}) - b(X_s^{(n-1)})) ds \right)^2 x \right] \leq \tau \int_0^t E [x^T (X_s^{(n)} - X_s^{(n-1)})^2 x] ds. \quad (9)$$

For the  $H$ -term, using the same idea as for the Uniqueness part:

$$\begin{aligned}
& H_t(X^{(n)}, X^{(n-1)}) \\
&= H_t(X^{(n)}, X^{(n-1)}) \pm \int_0^t g(X_s^{(n)}) dB_s f(X_s^{(n-1)}) \pm \int_0^t f(X_s^{(n-1)}) dB_s^T g(X_s^{(n)}) \\
&= \int_0^t g(X_s^{(n)}) dB_s (f(X_s^{(n)}) - f(X_s^{(n-1)})) + \int_0^t (f(X_s^{(n)}) - f(X_s^{(n-1)})) dB_s^T g(X_s^{(n)}) \\
&+ \int_0^t (g(X_s^{(n)}) - g(X_s^{(n-1)})) dB_s f(X_s^{(n)}) + \int_0^t f(X_s^{(n-1)}) dB_s^T (g(X_s^{(n)}) - g(X_s^{(n-1)})).
\end{aligned}$$

Using inequality (2) and iterating the same arguments (i.e. Lemma 6, Proposition 5, Lipschitz condition (4), boundedness condition (6)) as in the Uniqueness part we arrive at

$$E[x^T (X_t^{(n+1)} - X_t^{(n)})^2 x] \leq c_\tau \frac{(\beta t)^n}{n!}, \quad (10)$$

where we also used (8) in the last iteration. Notice that we obtain the same inequality after incorporating (9).

Define now

$$D_n := \sup_{t \leq \tau} |x^T (X_t^{(n+1)} - X_t^{(n)}) x|.$$

We want to prove that

$$\sum_{n=1}^{\infty} P(D_n \geq 1/n^2) < \infty. \quad (11)$$

Using the Chebyshev inequality

$$\sum_{n=1}^{\infty} P(D_n \geq 1/n^2) \leq \sum_{n=1}^{\infty} n^4 E(D_n^2).$$

So, it suffices to show that

$$E(D_n^2) \leq c_\tau \frac{(\beta \tau)^n}{n!}.$$

However,

$$\begin{aligned}
D_n &= \sup_{t \leq \tau} \left| x^T \left( \int_0^t (b(X_s^{(n)}) - b(X_s^{(n-1)})) ds + H_t(X^{(n)}, X^{(n-1)}) \right) x \right| \\
&\leq \int_0^\tau |x^T (b(X_s^{(n)}) - b(X_s^{(n-1)})) x| ds + \sup_{t \leq \tau} |x^T H_t(X^{(n)}, X^{(n-1)}) x|.
\end{aligned}$$



Thus,

$$D_n^2 \leq 2 \underbrace{\left( \int_0^\tau |x^T((b(X_s^{(n)}) - b(X_s^{(n-1)})))x| ds \right)^2}_{A_n^2} + 2 \underbrace{\left( \sup_{t \leq \tau} |x^T H_t(X^{(n)}, X^{(n-1)})x| \right)^2}_{C_n^2}.$$

From the Cauchy-Schwarz inequality and using the inequality (3), we can produce

$$\begin{aligned} \left( \int_0^\tau |x^T((b(X_s^{(n)}) - b(X_s^{(n-1)})))x| ds \right)^2 &\leq \int_0^\tau 1 ds \int_0^\tau |x^T((b(X_s^{(n)}) - b(X_s^{(n-1)})))x|^2 ds \\ &\leq \tau \int_0^\tau x^T (b(X_s^{(n)}) - b(X_s^{(n-1)}))^2 x ds. \end{aligned}$$

Hence, for a constant  $c_\tau^{(1)}$  depending on  $\tau$ ,

$$E(A_n^2) \leq c_\tau^{(1)} \frac{(\beta\tau)^n}{n!}.$$

For the  $H_t$ -term, since  $x^T H_t(X^{(n)} - X^{(n-1)})x$  is a martingale, by Doob's inequality

$$E(C_n^2) \leq 4E[x^T (H_\tau(X^{(n)}, X^{(n-1)}))^2 x].$$

Upon the same argument as for (10),

$$E(C_n^2) \leq c_\tau^{(2)} \frac{(\beta\tau)^n}{n!},$$

for some finite constant  $c_\tau^{(2)}$ , therefore, for  $c_\tau := \max(c_\tau^{(1)}, c_\tau^{(2)})$ ,

$$E(D_n^2) \leq c_\tau \frac{(\beta\tau)^n}{n!}.$$

Since (11) holds, by the Borel-Cantelli Lemma,

$$P(\liminf_n \{D_n < 1/n^2\}) = 1.$$

This says that

for almost all  $\omega \in \Omega$  there exists  $N(\omega)$  such that  $D_n < 1/n^2$  for all  $n \geq N(\omega)$ ,

which implies that  $\{X_t^{(n)}\}$  is a Cauchy sequence a.s., since

$$x^T X_t^{(n)} x = \sum_{i=0}^{n-1} x^T \left( X_t^{(i+1)} - X_t^{(i)} \right) x.$$

The conclusion is that there exists a process  $X_t$  with

$$X_t^{(n)} \xrightarrow{a.s.} X_t \text{ uniformly in } [0, \tau], \text{ as } n \rightarrow \infty.$$

It remains to prove that  $X_t$  satisfies (5), which is achieved from the inequality

$$E \left[ x^T \left( \int_0^t (b(X_s^{(n)}) - b(X_s)) ds + H_t(X^{(n)}, X) \right)^2 x \right] \leq c_\tau \beta \int_0^t E[x^T (X_s^{(n)} - X_s)^2 x] ds,$$

and taking  $n \rightarrow \infty$ . However, previous the inequality can be obtained repeating the same kind of arguments used along the proof. ■

## 4 Appendix

In this section we prove Proposition 5 and Lemma 6. In what follows,  $\|A\|$  represents the operator norm and  $\|A\|_F$  the Frobenius norm of a matrix  $A$ . It is well known that the norms in the space of matrices are equivalent because it is of finite dimension.

We now give some remarks regarding the matrix stochastic integral  $I := \int_0^t A_s dB_s C_s$  that we use below. Notice first that  $I$  is defined as the matrix where each  $(i, j)$  entry is given by

$$I(i, j) := \sum_{k=1}^d \sum_{r=1}^d \int_0^t A_s(i, k) C_s(r, j) dB_s(k, r),$$

where  $A(i, k)$ ,  $C(r, j)$  and  $B(k, r)$  are the corresponding entries of  $A$ ,  $C$  and  $B$ . Therefore, the existence of  $I$  occurs if

$$r \|A(i, k) C(r, j)\|_2^2 := E \left[ \int_0^t (A_s(i, k) C_s(r, j))^2 ds \right] < \infty,$$

for all  $i, k, r, j \in \{1, \dots, d\}$ . The above expression  $\|\bullet\|_2$  is a norm in a space of stochastic processes.

Let us see how we can construct a sequence of matrix step processes  $A^{(n)}$  and  $C^{(n)}$  such that  $E[\|I^{(n)} - I\|^2] \rightarrow 0$ ,  $n \rightarrow \infty$ , where  $I^{(n)} := \int_0^t A_s^{(n)} dB_s C_s^{(n)}$ . Take precisely the two sequences such that

$$\|A^{(n)}(i, k) C^{(n)}(r, j) - A(i, k) C(r, j)\|_2 \rightarrow 0$$

as  $n \rightarrow \infty$  for all  $i, k, r, j \in \{1, \dots, d\}$ . By the construction of the stochastic integral in one dimension, we can bound to have that  $E[(I^{(n)}(i, j) - I(i, j))^2] \rightarrow 0$  as  $n \rightarrow \infty$ , which helps to see that  $E[\|I^{(n)} - I\|_F^2] \rightarrow 0$ . Nevertheless, by the equivalence of norms  $E[\|I^{(n)} - I\|^2] \rightarrow 0$ .

## 4.1 Proof of Proposition 5

**i) For step processes.** First of all, we can check the formula for matrix step processes. In this case we have

$$E \left[ y^T \left( \int_0^{\cdot} A_s dB_s C_s \right)^2 x \right] = E \left[ y^T \left( \sum_{k=0}^{n-1} A_{s_k} (B_{s_{k+1}} - B_{s_k}) C_{s_k} \right)^2 x \right].$$

Notice that when expanding the square and taking expectation, the cross terms are vanished, then we have

$$E \left[ y^T \left( \int_0^{\cdot} A_s dB_s C_s \right)^2 x \right] = \sum_{k=0}^{n-1} E \left[ y^T (A_{s_k} (B_{s_{k+1}} - B_{s_k}) C_{s_k})^2 x \right].$$

Thus, we have to analyze

$$E \left[ y^T A_{s_k} (B_{s_{k+1}} - B_{s_k}) C_{s_k} A_{s_k} (B_{s_{k+1}} - B_{s_k}) C_{s_k} x \right],$$

written in a compact form as

$$E [a^T \beta c \beta b],$$

using the notation

$$a^T := y^T A_{s_k}, \quad \beta := B_{s_{k+1}} - B_{s_k}, \quad c := C_{s_k} A_{s_k}, \quad b := C_{s_k} x.$$

Notice that  $\beta$  is a matrix of independent normal r.v.s with mean 0 and variance  $s_{k+1} - s_k$ .

It will be easy to deduce the formula by analyzing the 2-dimensional case:

$$\begin{aligned} E [a^T \beta c \beta b] &= E \left[ (a_1 \ a_2) \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix} \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right] \\ &= E \left[ \begin{pmatrix} a_1 \beta_{11} + a_2 \beta_{21} \\ a_1 \beta_{12} + a_2 \beta_{22} \end{pmatrix}^T \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} b_1 \beta_{11} + b_2 \beta_{12} \\ b_1 \beta_{21} + b_2 \beta_{22} \end{pmatrix} \right] \\ &= E [a_1 c_{11} b_1 + a_1 c_{21} b_2 + a_2 c_{12} b_1 + a_2 c_{22} b_2] (s_{k+1} - s_k) \\ &= E \left[ (b_1 \ b_2) \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \right] (s_{k+1} - s_k) \\ &= E [b^T c a] (s_{k+1} - s_k) = E (x^T C^T C A A^T y) (s_{k+1} - s_k). \end{aligned}$$

This helps to see how the formula arises for step processes.

**ii) For more general processes.**

Let  $A$  and  $C$  be matrix stochastic processes where the stochastic integral  $I$  is well defined. Therefore, as mentioned above, there are approximating step processes  $A^{(n)}$  and  $C^{(n)}$  whose stochastic integral  $I^{(n)}$  converges to  $I$  in the  $L_2$ -norm.

By point **i)** above,

$$E [y^T (I^{(n)})^2 x] = \int_0^t E [x^T (C_s^{(n)})^T C_s^{(n)} A_s^{(n)} (A_s^{(n)})^T y] ds$$

for every  $n \geq 1$ . Then, we want to prove that

$$|E [y^T (I^{(n)})^2 x] - E [y^T I^2 x]| \rightarrow 0, \quad n \rightarrow \infty, \quad (12)$$

and that

$$\int_0^t E [x^T (C_s^{(n)})^T C_s^{(n)} A_s^{(n)} (A_s^{(n)})^T y] ds \rightarrow \int_0^t E [x^T C_s^T C_s A_s A_s^T y] ds, \quad n \rightarrow \infty. \quad (13)$$

For (12) we have

$$\begin{aligned} |E [y^T (I(n)^2 - I^2)x]| &= E [|y^T ((I(n) - I)I(n) + I(I(n) - I))x|] \\ &\leq E [|y^T (I(n) - I)I(n)x|] + E [|y^T I(I(n) - I)x|] \\ &\leq \|x\| \|y\| (E[\|I(n) - I\| \|I(n)\|] + E[\|I\| \|I(n) - I\|]) \\ &\leq \|x\| \|y\| \left( \sqrt{E[\|I(n)\|^2]} + \sqrt{E[\|I\|^2]} \right) \sqrt{E[\|I(n) - I\|^2]}. \end{aligned}$$

Hence, we obtain (12), because  $E[\|I(n) - I\|^2] \rightarrow 0$  as  $n \rightarrow \infty$ .

For (13), we need to calculate

$$\int_0^t E [x^T ((C_s^{(n)})^T C_s^{(n)} A_s^{(n)} (A_s^{(n)})^T - C_s^T C_s A_s A_s^T) y] ds.$$

Observe that we need to calculate

$$E \left[ \int_0^t (a_1^n(s) a_2^n(s) c_1^n(s) c_2^n(s) - a_1(s) a_2(s) c_1(s) c_2(s)) ds \right],$$

where  $a_1^n(s)$  and  $a_2^n(s)$  are arbitrary entries of  $A_s^{(n)}$ ,  $c_1^n(s)$  and  $c_2^n(s)$  of  $C_s^{(n)}$ , and similarly without the the index  $n$ , i.e.  $a_1(s)$  represents an entry of  $A_s$ .

After adding and subtracting  $a_1(s)a_2^n(s)c_1(s)c_2^n(s)$  we can split into two terms. Let us elaborate one of them, the other one is similar. We have that

$$\begin{aligned}
& E \left[ \int_0^t (a_1^n(s)a_2^n(s)c_1^n(s)c_2^n(s) - a_1(s)a_2^n(s)c_1(s)c_2^n(s))ds \right] \\
&= E \left[ \int_0^t (a_1^n(s)c_1^n(s) - a_1(s)c_1(s))a_2^n(s)c_2^n(s)ds \right] \\
&\leq E \left[ \sqrt{\int_0^t (a_1^n(s)c_1^n(s) - a_1(s)c_1(s))^2 ds} \sqrt{\int_0^t (a_2^n(s)c_2^n(s))^2 ds} \right] \\
&\leq \sqrt{E \left[ \int_0^t (a_1^n(s)c_1^n(s) - a_1(s)c_1(s))^2 ds \right] E \left[ \int_0^t (a_2^n(s)c_2^n(s))^2 ds \right]}.
\end{aligned}$$

where we used the Cauchy-Schwarz inequality twice, one for the integral and another one for the expectation. Since  $E[\int_0^t (a_1^n(s)c_1^n(s) - a_1(s)c_1(s))^2 ds]$  vanishes as  $n \rightarrow \infty$ , we obtain (13), and therefore the result.

## 4.2 Proof of Lemma 6

Define  $M_t := \int_0^t A_s dB_s C_s$ . Since  $M_t + M_t^T$  is symmetric,  $(M_t + M_t^T)^2$  is positive semidefinite, that is

$$0 \leq (M_t + M_t^T)^2 = M_t^2 + (M_t^T)^2 + M_t M_t^T + M_t^T M_t,$$

then

$$-(M_t M_t^T + M_t^T M_t) \leq M_t^2 + (M_t^T)^2.$$

So that

$$\begin{aligned}
-E[x^T(M_t M_t^T + M_t^T M_t)x] &\leq E[x^T M_t^2 x] + E[x^T (M_t^T)^2 x] \\
&\leq |E[x^T M_t^2 x]| + |E[x^T (M_t^T)^2 x]|.
\end{aligned}$$

Now, for each  $t \in [0, \tau]$ , we can find  $\alpha_t > 0$  such that

$$\alpha_t |E[x^T(M_t M_t^T + M_t^T M_t)x]| \leq |E[x^T M_t^2 x]| + |E[x^T (M_t^T)^2 x]|.$$

From the continuous trajectories of  $B_t$ , we have that  $\alpha : [0, \tau] \rightarrow (0, \infty)$  is actually continuous. Let us then define  $\delta := \min_{t \in [0, \tau]} \alpha_t$ . Then

$$\begin{aligned} E [x^T (M_t + M_t^T)^2 x] &= E[x^T M_t^2 x] + E[x^T (M_t^T)^2 x] \\ &\quad + E[x^T (M_t M_t^T + M_t^T M_t) x] \\ &\leq |E[x^T M_t^2 x]| + |E[x^T (M_t^T)^2 x]| \\ &\quad + \frac{1}{\delta} \{ |E[x^T M_t^2 x]| + |E[x^T (M_t^T)^2 x]| \}. \end{aligned}$$

Defining  $\beta := 1 + \delta^{-1}$  gives the inequality.

## References

- [1] Bru, M.-F. (1989). Diffusions of Perturbed Principal Component Analysis. *Journal of Multivariate Analysis* **29**, pp. 127–136.
- [2] Bru, M.-F. (1991). Wishart Processes. *Journal of Theoretical Probability* **4**(4), pp. 725–751.
- [3] Cuchiero, C., Filipović, D., Mayerhofer, E., Teichmann, J. (2011). Affine processes on positive semidefinite matrices. *The Annals of Applied Probability* **21**(2), pp. 397–463.
- [4] Donati-Martin, C., Doumerc, Y., Matsumoto, H. and Yor, M. (2004). *Some properties of the Wishart processes and a matrix extension of the Hartman-Watson laws*. Publ. Res. Inst. Math. Sci. **40**(4), pp. 1385–1412
- [5] Graczyk, P. and Malecki, J. (2013). Multidimensional Yamada-Watanabe theorem and its applications to particle systems. *Journal of Mathematical Physics* **54**(2), pp. 725–751.
- [6] Graczyk, P. and Mayerhofer, E. (2011). Stochastic analysis Methods in Wishart Theory. *CIMPA Workshop*.
- [7] Horn, R.A. and Johnson, C.R. (1991). *Topics in Matrix Analysis*, Cambridge University Press.
- [8] Ikeda, N. and Watanabe, S. (1981). *Stochastic Differential Equations and Diffusion Processes*, North-Holland Publishing Company.
- [9] Kang, W. and Kang, C. (2013). Transform formulae for linear functionals of affine processes and their bridges on positive semidefinite matrices. *Stochastic Processes and their Applications* **13**, pp. 2419–2445.

- [10] Katori, M. and Tanemura, H. (2004). Symmetry of matrix-valued stochastic processes and noncolliding diffusion particle systems. *Journal of Mathematical Physics* **45**(8), pp. 3058–3085.
- [11] McKean, H.P. (1969). *Stochastic Integrals*, Academic Press.
- [12] König, W. and O’Connell, N. (2001). Eigenvalues of the Laguerre process as non-colliding squared Bessel processes. *Electronic Communications in Probability* **6**, pp. 107–114.
- [13] Mayerhofer, E., Pfaffel, O. and Stelzer, R. (2011). On strong solutions for positive definite jump diffusions. *Stochastic Processes and their Applications* **121**, pp. 2072–2086.
- [14] Pfaffel, O. (2008). *Wishart Processes*, Ph.D. thesis.
- [15] Rogers, L.C.G. and Williams, D. (2000). *Diffusions, Markov Processes and Martingales*, Cambridge University Press.
- [16] Stelzer, R.J. (2007). *Multivariate Continuous Time Stochastic Volatility Models Driven by a Lévy Process*, Ph.D. thesis.
- [17] Trujillo Rivera, E.A. (2011). *El Proceso de Wishart y la Dinámica de sus Eigenvalores y Propiedades Distribucionales Vía Cálculo Estocástico*, M.Sc. thesis.
- [18] Tudor, C. (2002). *Procesos Estocásticos*, Aportaciones Matemáticas, Sociedad Matemática Mexicana.
- [19] Zhang, F. (2011). *Matrix Theory: Basic Results and Techniques* (2nd ed.), Springer.